

Linear spanning sets for matrix spaces

G. Micheli^{a,1}, J. Rosenthal^{a,1}, P. Vettori^{b,1,2,*}

^aUniversity of Zurich, Winterthurststrasse 190, CH-8057 Zürich, Switzerland

^bUniversity of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

Abstract

Necessary and sufficient conditions are given on matrices A , B and S , having entries in some field \mathbb{F} and suitable dimensions, such that the linear span of the terms $A^i S B^j$ over \mathbb{F} is equal to the whole matrix space.

This result is then used to determine the cardinality of subsets of $\mathbb{F}[A]S\mathbb{F}[B]$ when \mathbb{F} is a finite field.

Keywords: Matrices, linear span, cyclic matrices, finite fields.

2010 MSC: 15A03, 15A69

1. Introduction

We start by stating a purely linear algebra problem:

Problem 1.1. *Let m, n be integers and \mathbb{F} be any field. Let A, S, B be matrices having entries in \mathbb{F} of dimensions $m \times m$, $m \times n$ and $n \times n$ respectively. Give necessary and sufficient conditions for the \mathbb{F} -linear span of $\{A^i S B^j\}_{i,j \in \mathbb{N}}$ to be equal to the whole matrix space $\mathbb{F}^{m \times n}$.*

*Corresponding author

Email addresses: giacomo.micheli at math.uzh.ch (G. Micheli), rosenthal at math.uzh.ch (J. Rosenthal), pvettori at ua.pt (P. Vettori)

¹Authors supported in part by Swiss National Science Foundation grant SNF no. 149716.

²This work was supported by Portuguese funds through the CIDMA (Center for Research and Development in Mathematics and Applications) and the Portuguese Foundation for Science and Technology (“FCT-Fundação para a Ciência e a Tecnologia”), within project PEst-OE/MAT/UI4106/2014.

A solution to this problem will be provided in Section 3.

Starting with Section 4 we will assume that the base field \mathbb{F} represents the finite field $\mathbb{F} = \mathbb{F}_q$ having cardinality q . Under these conditions and the conditions that $\gcd(m, n) = 1$ and the characteristic polynomials of the matrices A and B are irreducible we are able to show in Section 4 that $\{A^i S B^j\}_{i,j \in \mathbb{N}}$ spans the whole vector space $\mathbb{F}^{m \times n}$ as soon as $S \neq 0$.

In Section 5 we will prove that whenever the set $\{A^i S B^j\}_{i,j \in \mathbb{N}}$ spans the whole matrix ring as a vector space over the finite field \mathbb{F} , we are able to explicitly compute the cardinality $|\mathbb{F}[A]S\mathbb{F}[B]|$. A particular instance of this computation (i.e. when S is the identity matrix and A, B have irreducible characteristic polynomial) has already been approached via inequalities in [1].

2. Notation and Preliminaries

Let \mathbb{F} be a field and denote by $\langle S \rangle_{\mathbb{F}}$ the linear span over \mathbb{F} of a set S of elements in some \mathbb{F} -vector space. Entries, rows and columns of matrices are indexed by integers starting from zero; I_n and, respectively, $0_{m \times n}$ denote the $n \times n$ identity matrix and the $m \times n$ zero matrix — indices may be omitted when no ambiguity arises.

Moreover, given $M \in \mathbb{F}^{n \times n}$,

- the **minimal polynomial** μ_M of M is the monic generator of the ideal $\{p(s) \in \mathbb{F}[s] : p(M) = 0\}$;
- the **characteristic polynomial** of M is $\chi_M(s) = \det(sI - M)$;
- \mathcal{E}_M is the set of eigenvalues of M , i.e., the zeros of χ_M in some field extension of \mathbb{F} ;
- \mathcal{L}_M^λ and \mathcal{R}_M^λ are the left and, respectively, right eigenspaces of M associated with $\lambda \in \mathcal{E}_M$;
- $\mathcal{L}_M = \bigcup_{\lambda \in \mathcal{E}_M} \mathcal{L}_M^\lambda \setminus \{0\}$ and $\mathcal{R}_M = \bigcup_{\lambda \in \mathcal{E}_M} \mathcal{R}_M^\lambda \setminus \{0\}$ are the sets of left and, respectively, right eigenvectors of M .
- M is **cyclic** (or non-derogatory) if one of the following equivalent conditions holds true:
 - $\mu_M = \chi_M$;
 - M is similar to a companion matrix;

- each eigenspace of M has dimension 1, i.e., every eigenvector has geometric multiplicity 1.

The definition of the Kronecker product and some of its properties are given next. More details may be found in [2, Section 12.1].

Definition 2.1. The **Kronecker product** of matrices $M \in \mathbb{F}^{m \times p}$ and $N \in \mathbb{F}^{n \times q}$ is the block matrix

$$M \otimes N = [m_{i,j}N]_{0 \leq i < m, 0 \leq j < p} \in \mathbb{F}^{mn \times pq},$$

representing the tensor product of the linear maps corresponding to M and N . Therefore, it satisfies the property

$$(M \otimes N)(P \otimes Q) = MP \otimes NQ, \quad (1)$$

whenever the matrix products on the right side can be computed.

The **(column) vectorization** of M is the (column) vector $\mathbf{v}(M) \in \mathbb{F}^{mp}$ formed by stacking the columns of M . Note that $\mathbf{v} : \mathbb{F}^{m \times p} \rightarrow \mathbb{F}^{mp}$ is an isomorphism of \mathbb{F} -vector spaces, establishing a correspondence between entry (i, j) of M and entry $i + mj$ of $\mathbf{v}(M)$.

Using this notation, given three matrices M, X, N of suitable dimensions,

$$\mathbf{v}(MXN) = (N^\top \otimes M) \mathbf{v}(X). \quad (2)$$

3. A basis for the vector space of $m \times n$ matrices

Let matrices A, B , and S as in Problem 1.1 and define

$$\mathcal{V}_{A,B;S} = \langle \{A^i S B^j\}_{i,j \geq 0} \rangle_{\mathbb{F}}.$$

In this and in the following section, conditions will be given that ensure that the dimension of $\mathcal{V}_{A,B;S}$ is maximal, i.e., equal to mn .

Theorem 3.1. *Let $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, and $S \in \mathbb{F}^{m \times n}$ and consider the following conditions:*

$$\mathcal{V}_{A,B;S} = \mathbb{F}^{m \times n}; \quad (3)$$

$$A \text{ and } B \text{ are cyclic}; \quad (4)$$

$$uSv \neq 0, \forall u \in \mathcal{L}_A, v \in \mathcal{R}_B. \quad (5)$$

Then, (3) \Leftrightarrow ((4) and (5)).

Remark 3.2. The previous theorem has also an impact in Cryptography since it gives necessary and sufficient conditions for the attack in [3, Section 3] to be performed in *provable* polynomial time.

Before proving the theorem, two lemmas will be stated. The first one provides a logical equivalence, which will be used within different proofs.

Lemma 3.3. *Given three conditions A , B , and C , then $A \Leftrightarrow (B \text{ and } C)$ is equivalent to: $(A \Rightarrow B)$ and $(B \Rightarrow (A \Leftrightarrow C))$.*

PROOF. It is easy to check that both conditions are equivalent to the negation of A , when B is false, and to $A \Leftrightarrow C$, when B is true. \square

The second lemma is well known (see [4, 5]) in the case $\mathbb{F} = \mathbb{C}$. For completeness, a self-contained proof will be given here.

Lemma 3.4. *Let $H \in \mathbb{F}^{p \times p}$, $K \in \mathbb{F}^{p \times q}$ and assume that $\mathcal{E}_H \subseteq \mathbb{E}$, extension field of \mathbb{F} . Then, for any $d \geq \deg \mu_H$.*

$$\text{rank}_{\mathbb{F}} \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} = p \Leftrightarrow \text{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - H & K \end{bmatrix} = p, \forall \lambda \in \mathcal{E}_H.$$

PROOF. Observe that for any matrix M with entries in \mathbb{F} , $\text{rank}_{\mathbb{F}} M = \text{rank}_{\mathbb{E}} M$, since the rank depends only on the invertibility (in \mathbb{F}) of square submatrices of M . So, this equivalent statement will be proved:

$$\text{rank}_{\mathbb{E}} \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} < p \Leftrightarrow \exists \lambda \in \mathcal{E}_H : \text{rank}_{\mathbb{E}} \begin{bmatrix} sI - H & K \end{bmatrix} < p.$$

“ \Rightarrow ”: Be $u \in \mathbb{E}^{1 \times p}$ a nonzero vector such that $u \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} = 0$ and be $a \in \mathbb{E}[s]$ any generator of the principal ideal $\mathcal{I} = \{f \in \mathbb{E}[s] : uf(H) = 0\}$. Since $\mu_H \in \mathcal{I}$, $\deg a \leq \deg \mu_H \leq d$ and $a(\lambda) = 0$ for some $\lambda \in \mathcal{E}_H$. Write $a(s) = (\lambda - s)b(s)$, being $b(s) = \sum_{i=0}^{d-1} b_i s^i \notin \mathcal{I}$. Hence, $v = ub(H) \neq 0$. Moreover,

$$vK = ub(H)K = \sum_{i=0}^{d-1} b_i u H^i K = \sum_{i=0}^{d-1} b_i 0 = 0$$

and $0 = ua(H) = u(\lambda I - H)b(H) = v(\lambda I - H)$. Thus, $v \begin{bmatrix} \lambda I - H & K \end{bmatrix} = 0$.

“ \Leftarrow ”: There exist $\lambda \in \mathcal{E}_H$ and a nonzero $u \in \mathbb{E}^{1 \times p}$ such that $u \begin{bmatrix} \lambda I - H & K \end{bmatrix} = 0$, i.e., $uH = \lambda u$ and $uK = 0$. Hence,

$$u \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} = u \begin{bmatrix} K & \lambda K & \cdots & \lambda^{d-1}K \end{bmatrix} = 0.$$

\square

PROOF (OF THEOREM 3.1). Consider the new conditions (4a): A is cyclic and (4b): B is cyclic, so that (4) is equivalent to (4a) and (4b). Therefore, the equivalence (3) \Leftrightarrow ((4a) and (4b) and (5)) will be proved.

First of all, note that matrices $\{A^i SB^j\}$ generate $\mathbb{F}^{m \times n}$ if and only if the corresponding vectors $\{\mathbf{v}(A^i SB^j)\}$ generate \mathbb{F}^{mn} . Therefore, we get that

$$(3) \Leftrightarrow \langle \{\mathbf{v}(A^i SB^j)\}_{i,j \geq 0} \rangle_{\mathbb{F}} = \mathbb{F}^{mn}. \quad (6)$$

By (2) and (1), it follows that

$$\mathbf{v}(A^i SB^j) = \mathbf{v}(A^i SB^j I_n) = (I_n \otimes A^i) \mathbf{v}(SB^j) = (I_n \otimes A)^i \mathbf{v}(SB^j).$$

Let $F = I_n \otimes A \in \mathbb{F}^{mn \times mn}$, which is a block diagonal matrix, and be G the $mn \times n$ matrix whose columns are $\mathbf{v}(SB^j)$, $0 \leq j < n$. The (right) image of G , i.e., its column span, corresponds through \mathbf{v} to the span of SB^j , $0 \leq j < n$. Analogously, for any $0 \leq i < m$, the image of $F^i G$ corresponds to the span of $A^i SB^j$, $0 \leq j < n$. Hence, by the Cayley-Hamilton Theorem,

$$\langle \{\mathbf{v}(A^i SB^j)\}_{i,j \geq 0} \rangle_{\mathbb{F}} = \text{img}_{\mathbb{F}} [G \ FG \ \cdots \ F^{m-1}G]. \quad (7)$$

Observe that the degree of the minimal polynomial $\mu_F = \mu_{I \otimes A} = \mu_A$ cannot be greater than m and so, by (6), (7) and Lemma 3.4, we can state that

$$\begin{aligned} (3) &\Leftrightarrow \text{img}_{\mathbb{F}} [G \ FG \ \cdots \ F^{m-1}G] = \mathbb{F}^{mn} \\ &\Leftrightarrow \text{rank}_{\mathbb{F}} [\lambda I - F \ G] = mn, \ \forall \lambda \in \mathcal{E}_A, \end{aligned} \quad (8)$$

being \mathbb{F} the extension field of \mathbb{F} containing the eigenvalues of F , i.e., of A .

In order to determine the conditions that guarantee that the rank of the polynomial matrix $C(s) = [sI - F \ G]$ does not drop as $s \in \mathcal{E}_A$, it is necessary to analyze the structure of $C(s)$ with greater detail.

Denote by G_i , $0 \leq i < n$, the $m \times n$ blocks forming matrix G . Then

$$C(s) = [sI - F \ G] = \begin{bmatrix} sI - A & & & G_0 \\ & sI - A & & G_1 \\ & & \ddots & \vdots \\ & & & sI - A & G_{n-1} \end{bmatrix}. \quad (9)$$

Now, let α be any eigenvalue of A with geometric multiplicity h and observe that the rank of the block-diagonal matrix $\alpha I - F$ (the first mn columns of $C(\alpha)$) is

equal to $n(m - h) = mn - nh$. Since matrix G has n columns, the rank of $C(\alpha)$ cannot exceed $mn - nh + n = mn - n(h - 1)$. This shows that for condition (8) to hold, it is necessary to have $h = 1$, i.e., A must be cyclic — by equivalence (8), this shows that (3) \Rightarrow (4a).

On the other hand, by assuming that A is cyclic, it follows that the rank of $C(\alpha)$ is mn if and only if for every $w \neq 0$ such that $w(\alpha I - F) = 0$, we have that $wC(\alpha) \neq 0$. Since $\alpha I - F = I_n \otimes (\alpha I - A)$, it turns out that $w(\alpha I - F) = 0$ if and only if $w = [u_0 \ u_1 \ \cdots \ u_{n-1}]$, with $u_i \in \mathcal{L}_A^\alpha$, $0 \leq i < n$. Therefore,

$$\begin{aligned} wC(\alpha) &= [u_0 \ u_1 \ \cdots \ u_{n-1}] \begin{bmatrix} \alpha I - A & & & G_0 \\ & \alpha I - A & & G_1 \\ & & \ddots & \vdots \\ & & & \alpha I - A & G_{n-1} \end{bmatrix} \\ &= [0 \ u_0 G_0 + u_1 G_1 + \cdots + u_{n-1} G_{n-1}] = [0 \ g], \ g \in \mathbb{E}^{1 \times n}. \end{aligned} \quad (10)$$

Since the eigenspace \mathcal{L}_A^α has dimension 1, is it generated by one (eigen)vector, say $u \neq 0$, whence $u_i = \gamma_i u$, $\gamma_i \in \mathbb{E}$ for $0 \leq i < n$, not all zero. This means that

$$g = \gamma_0 u G_0 + \gamma_1 u G_1 + \cdots + \gamma_{n-1} u G_{n-1}$$

is not zero if and only if vectors $\{u G_i\}_{0 \leq i < n}$ are linearly independent. Hence, by equivalence (8), condition (4a) implies that (3) is equivalent to the linear independence of $\{u G_i\}_{0 \leq i < n}$, for every $u \in \mathcal{L}_A$. We already proved that (3) \Rightarrow (4a) and so, by Lemma 3.3, it follows that

$$(3) \Leftrightarrow ((4a) \text{ and } \forall u \in \mathcal{L}_A, \{u G_i\}_{0 \leq i < n} \text{ are } \mathbb{E}\text{-linearly independent}). \quad (11)$$

Consider now any $u \in \mathbb{E}^{1 \times m}$ and define the matrix

$$D = (I_n \otimes u)G = \begin{bmatrix} uG_0 \\ uG_1 \\ \vdots \\ uG_{n-1} \end{bmatrix} \in \mathbb{E}^{n \times n}.$$

Moreover, for every $0 \leq i < n$ and $0 \leq j < n$, let $(SB^j)_i$ be the i -th column of SB^j .

By definition, the j -th column of G is $\vee(SB^j)$, which contains, stacked, vectors $(SB^j)_i$. Therefore, in particular, the j -th column of G_i , is $(SB^j)_i$. Consequently,

the j -th component of uG_i , which is the entry at (i, j) of D , is $u(SB^j)_i$. At the same time, this value is the i -th component (column) of uSB^j , i.e., the entry at (j, i) of the matrix whose rows are uSB^j . In other words,

$$D^\top = \begin{bmatrix} uSB^0 \\ uSB^1 \\ \vdots \\ uSB^{n-1} \end{bmatrix}.$$

Since D is square, its rows are linearly independent if and only if its columns share the same property. Applying again Lemma 3.4 with $H = B^\top$ and $K = (uS)^\top$, we get that

$$\begin{aligned} \{uG_i : 0 \leq i < n\} \text{ are } \mathbb{E}\text{-linearly independent} & \Leftrightarrow (12) \\ \{uSB^j : 0 \leq j < n\} \text{ are } \mathbb{E}\text{-linearly independent} & \Leftrightarrow \\ \text{rank}_{\mathbb{E}} \begin{bmatrix} (uS)^\top & B^\top(uS)^\top & \cdots & (B^\top)^{n-1}(uS)^\top \end{bmatrix} = n & \Leftrightarrow \\ \text{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - B^\top & (uS)^\top \end{bmatrix} = \text{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - B \\ uS \end{bmatrix} = n, \forall \lambda \in \mathcal{E}_B. & \end{aligned}$$

As before, consider $E(s) = \begin{bmatrix} sI - B \\ uS \end{bmatrix} \in \mathbb{E}^{(n+1) \times n}[s]$ and any $\beta \in \mathcal{E}_B$. Since $\beta I - B$ has rank $n - k$, where k is the geometric multiplicity of β , $\text{rank}_{\mathbb{E}} E(\beta) \leq n - k + 1$. We conclude that, when (12) holds, then $k = 1$, i.e., (12) \Rightarrow (4b), i.e., B is cyclic.

By assuming that B is cyclic, the rank of $E(\beta)$ is effectively n if $E(\beta)v \neq 0$ for any $v \in \mathcal{R}_B^\beta$. Since $(\beta I - B)v = 0$, condition $E(\beta)v \neq 0$ reduces to $uSv \neq 0$:

if B is cyclic, i.e., (4b) holds, (12) $\Leftrightarrow uSv \neq 0, \forall v \in \mathcal{R}_B$.

Thus, by Lemma 3.3, (12) \Leftrightarrow ((4b) and $uSv \neq 0, \forall v \in \mathcal{R}_B$). This, together with (11), concludes the proof. \square

Example 3.5. Consider the following matrices, with $m, n \geq 2$:

$$A = \begin{bmatrix} 0 & 0 \\ I_{m-1} & 0 \end{bmatrix} \in \mathbb{F}^{m \times m}, \quad B = \begin{bmatrix} 0 & I_{n-1} \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{n \times n}, \quad S = \begin{bmatrix} 1 & 0 \\ 0 & 0_{(m-1) \times (n-1)} \end{bmatrix} \in \mathbb{F}^{m \times n}.$$

Both A and B are already in (left and right, respectively) Jordan canonical form. Therefore, their only eigenvalue is $\lambda = 0$, they are nilpotent and cyclic with minimal polynomials $\mu_A(s) = s^m$ and $\mu_B(s) = s^n$, and their eigenspaces are generated by $u = [1 \ 0 \ \cdots \ 0]$ (left eigenvector of A) and $v = [1 \ 0 \ \cdots \ 0]^\top$ (right eigenvector of B).

Even though S has rank 1, $uSv = 1 \neq 0$, whence conditions (4) and (5) of Theorem 3.1 are satisfied. Therefore, \mathbb{F} -linear combinations of matrices $E_{i,j} = A^i S B^j$, with $0 \leq i < m$ and $0 \leq j < n$, generate $\mathbb{F}^{m \times n}$ for any field \mathbb{F} .

Indeed, it is straightforward to check that each $E_{i,j}$ is one of the mn elements of the canonical basis of $\mathbb{F}^{m \times n}$, having its unique nonzero entry, equal to 1, at position (i, j) . In other words, $v(E_{i,j})$ is the $i + mj$ -th vector of the canonical basis of \mathbb{F}^{mn} .

To the authors' knowledge, equality (3) and the kind of equivalent conditions that were presented in Theorem 3.1 have not been considered in the literature before (not even when $m = n$: see, for instance, the survey [6] containing a small section about spanning sets of matrix algebras).

A comparison with previous results can be made only in the case $m = n = 2$, verifying that $\mathbb{F}^{2 \times 2}$ is spanned by linear combinations of $A^i B^j$, $i, j = 0, 1$, if and only if it can be generated by A and B as a matrix algebra. (The well-known criterium for the latter problem, presented in the following proposition, can be found, for example, in [7], where it is thoroughly investigated.)

Proposition 3.6. *Let $A, B \in \mathbb{F}^{2 \times 2}$ and $S = I$. Then, the commutator $[A, B] = AB - BA$ is invertible if and only if conditions (4) and (5) hold.*

PROOF. Notice that adding a scalar matrix cI , $c \in \mathbb{F}$, to A or B does not change both the spanned space and the generated algebra, nor the commutator $[A, B]$. Therefore, we shall assume that A and B have zero trace.

First, observe that A is not cyclic if and only if its canonical Jordan form is a scalar matrix if and only if A itself is a scalar matrix, i.e., zero. Therefore, if either A or B is not cyclic, $[A, B] = 0$. This proves that

$$[A, B] \text{ is invertible} \Rightarrow (4). \quad (13)$$

Assume now (4), both A and B are cyclic, and suppose, without loss of generality, that A is in Jordan form. This means that

$$A = \begin{bmatrix} a & b \\ 0 & -a \end{bmatrix}, \quad B = \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix}, \quad \text{and} \quad [A, B] = \begin{bmatrix} b\gamma & 2a\beta - 2\alpha b \\ -2a\gamma & -b\gamma \end{bmatrix}. \quad (14)$$

In order to be cyclic, i.e., not zero, matrix B must satisfy $\alpha \neq 0$, $\beta \neq 0$ or $\gamma \neq 0$. For matrix A , the two following cases are possible.

1. $a = 0$ and $b = 1$: $\mathcal{L}_A \cup \{0\} = \mathcal{L}_A^0$ is generated by $u = \begin{bmatrix} 0 & 1 \end{bmatrix}$. If $\gamma = 0$ then $\alpha \in \mathcal{E}_B$ and $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathcal{R}_B^\alpha$, satisfying $uv = 0$. On the other hand, if

$uv = 0$, with $v \in \mathcal{R}_B^\lambda$ for some $\lambda \in \mathcal{E}_B$, then $v = \begin{bmatrix} x \\ 0 \end{bmatrix}$, $x \neq 0$. By definition,

$$Bv = \lambda v \Leftrightarrow \begin{bmatrix} \alpha x \\ \gamma x \end{bmatrix} = \begin{bmatrix} \lambda x \\ 0 \end{bmatrix} \Leftrightarrow \gamma = 0.$$

By (14), it easy to check that $[A, B]$ is singular if and only if $\gamma = 0$, thus proving that $[A, B]$ invertible \Leftrightarrow (5).

2. $a \neq 0$ and $b = 0$: both $u = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $u = \begin{bmatrix} 0 & 1 \end{bmatrix}$ belong to \mathcal{L}_A . If $\beta\gamma = 0$ then, similarly to the previous case, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ belong to \mathcal{R}_B , being possible to satisfy $uv = 0$ with a nonzero $v \in \mathcal{R}_B$. Vice versa, if $uv = 0$ for some $v = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathcal{R}_B$, then $xy = 0$. It turns out that $Bv = \lambda v$ implies that $x = 0 \Rightarrow \beta = 0$ and $y = 0 \Rightarrow \gamma = 0$, therefore $\beta\gamma = 0$.

Concluding, by (14), $[A, B]$ invertible $\Leftrightarrow \beta\gamma \neq 0 \Leftrightarrow$ (5).

We showed that, in both cases, when (4) holds, then $[A, B]$ is invertible \Leftrightarrow (5). The statement follows by (13) and Lemma 3.3. \square

When conditions (4) and (5) of Theorem 3.1 are not satisfied, matrices $A^i S B^j$, with $0 \leq i < m$ and $0 \leq j < n$, are linearly dependent. However, something more can be said about the dimension of the space they generate.

The general case demands an extremely complicated notation: only the case of cyclic and diagonalizable matrices A and B will be considered in this paper.

Theorem 3.7. *Let $S \in \mathbb{F}^{m \times n}$ and suppose that $A \in \mathbb{F}^{m \times m}$ and $B \in \mathbb{F}^{n \times n}$ are cyclic and diagonalizable. In particular, be $U \in \mathbb{F}^{m \times m}$ and $V \in \mathbb{F}^{n \times n}$ two invertible matrices, in some extension field \mathbb{E} of \mathbb{F} , such that $U A U^{-1}$ and $V^{-1} B V$ are diagonal.*

Then, the dimension of $\mathcal{V}_{A,B,S}$, is equal to the number of nonzero entries of $U S V$.

Before proving Theorem 3.7, we introduce the necessary notation and state a fundamental lemma.

Given $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, and $S \in \mathbb{F}^{m \times n}$, let $r_{i,j} = \mathfrak{v}(A^i S B^j)$ and define

$$R_{A,B,S} = \begin{bmatrix} r_{0,0} & r_{1,0} & \cdots & r_{m-1,0} & r_{0,1} & r_{1,1} & \cdots & r_{m-1,n-1} \end{bmatrix} \in \mathbb{F}^{mn \times mn}. \quad (15)$$

Then, given $v \in \mathbb{F}^n$, $\text{diag}(v) \in \mathbb{F}^{n \times n}$ is the diagonal matrix defined by the components of v . Moreover, let $\text{diag}(M) = \text{diag}(\mathfrak{v}(M))$ for any matrix M .

Finally, let $\bar{x}^n = \begin{bmatrix} 1 & x & \cdots & x^{n-1} \end{bmatrix}$ and be $\mathcal{V}_{x_1, \dots, x_k}^n$ the matrix whose rows are $\bar{x}_1^n, \dots, \bar{x}_k^n$.

Lemma 3.8. Let $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, and $S \in \mathbb{F}^{m \times n}$. Suppose that $u_h \in \mathcal{L}_A^{\alpha_h}$, $0 \leq h < s$, and $v_k \in \mathcal{L}_B^{\beta_k}$, $0 \leq k < t$, are the rows and, respectively, columns of matrices $U \in \mathbb{F}^{s \times m}$ and $V \in \mathbb{F}^{n \times t}$ in a suitable extension field \mathbb{E} of \mathbb{F} . Then,

$$(V^\top \otimes U)R_{A,B;S} = \text{diag}(USV)(\mathcal{V}_{\beta_1, \dots, \beta_t}^n \otimes \mathcal{V}_{\alpha_1, \dots, \alpha_s}^m). \quad (16)$$

PROOF. Observe that, for any row u_h of U and column v_k of V , there exist $\alpha_h \in \mathcal{E}_A$ and $\beta_k \in \mathcal{E}_B$ such that $u_h \in \mathcal{L}_A^{\alpha_h}$ and $v_k \in \mathcal{R}_B^{\beta_k}$. Thus,

$$(v_k^\top \otimes u_h) \mathfrak{v}(A^i S B^j) = u_h A^i S B^j v_k = u_h S v_k \alpha_h^i \beta_k^j$$

and, from (15), it follows that

$$(v_k^\top \otimes u_h)R_{A,B;S} = u_h S v_k (\bar{\beta}_k^n \otimes \bar{\alpha}_h^m).$$

Stacking up all these equalities, we get equation (16). \square

Remark 3.9. Using Lemma 3.8, implication (3) \Rightarrow (5) of Theorem 3.1 can be proved in a much simpler way.

Indeed, suppose that the nonzero left-eigenvector $u \in \mathcal{L}_A^\alpha$ and right-eigenvector $v \in \mathcal{R}_B^\beta$ satisfy $uSv = 0$. Then, taking $U = u$ and $V = v$ in formula (16), we get

$$(v^\top \otimes u)R_{A,B;S} = (uSv)(\bar{\beta}^n \otimes \bar{\alpha}^m) = 0,$$

showing that $R_{A,B;S}$ does not have full rank. Therefore, its columns $\mathfrak{v}(A^i S B^j)$ are linearly dependent and the set of matrices $A^i S B^j$ cannot generate $\mathbb{F}^{m \times n}$.

PROOF (OF THEOREM 3.7). Let α_h , $0 \leq h < m$ and β_k , $0 \leq k < n$, be the left eigenvalues of A associated with the rows of U and, respectively, the right eigenvalues of B associated with the columns of V .

Since A and B are cyclic and diagonalizable, they have no repeated eigenvalues, whence $\mathcal{V}_{\alpha_1, \dots, \alpha_s}^m$ and $\mathcal{V}_{\beta_1, \dots, \beta_n}^n$ are invertible Vandermonde matrices.

By Lemma 3.8, we have that

$$(V^\top \otimes U)R_{A,B;S} = \text{diag}(USV)(\mathcal{V}_{\beta_1, \dots, \beta_n}^n \otimes \mathcal{V}_{\alpha_1, \dots, \alpha_s}^m),$$

where both Kronecker products are invertible. So, $\text{rank } R_{A,B;S} = \text{rank } \text{diag}(USV)$, which is equal to the number of nonzero entries of USV .

Since by definition (16), the (column) rank of $R_{A,B;S}$ is equal to the dimension of the space spanned by $A^i S B^j$, the proof is concluded. \square

4. The irreducible case

For the remainder of the paper we will assume that $\mathbb{F} = \mathbb{F}_q$ represents the finite field of order q .

The main result of this section will provide a necessary and sufficient condition for matrices A, B having irreducible characteristic polynomial which guarantees that condition (3) of Theorem 3.1 holds true:

Theorem 4.1. *Let \mathbb{F} be a finite field, $A \in \mathbb{F}^{m \times m}$, $S \in \mathbb{F}^{m \times n}$ and $B \in \mathbb{F}^{n \times n}$. Suppose that A and B have irreducible characteristic polynomials. Then,*

$$\mathcal{V}_{A,B;S} = \mathbb{F}^{m \times n}, \forall S \neq 0 \text{ if and only if } \gcd(m, n) = 1.$$

PROOF. Define the \mathbb{F} -linear map

$$\begin{aligned} \psi : \mathbb{F}^{m \times n} &\rightarrow \mathbb{F}^{m \times n} \\ Z = [z_{i,j}] &\mapsto \sum_{\substack{0 \leq i < m \\ 0 \leq j < n}} z_{i,j} A^i S B^j \end{aligned} \quad (17)$$

and note that $\mathcal{V}_{A,B;S}$ is the image of ψ . Therefore, we need to prove that $\ker \psi = \{0\}, \forall S \neq 0 \Leftrightarrow \gcd(m, n) = 1$. By (2) we obtain that

$$\mathbb{v}(\psi(Z)) = \mathbb{v} \left(\sum_{\substack{0 \leq i < m \\ 0 \leq j < n}} z_{i,j} A^i S B^j \right) = \sum_{\substack{0 \leq i < m \\ 0 \leq j < n}} z_{i,j} (B^j)^\top \otimes A^i \mathbb{v}(S).$$

Hence, by injectivity of \mathbb{v} , it follows that ψ is injective (for any choice of $S \neq 0$) if and only if the kernel of matrix $M = \sum_{\substack{0 \leq i < m, 0 \leq j < n}} z_{i,j} (B^j)^\top \otimes A^i$ is trivial, i.e., M has no zero eigenvalues whenever $Z \neq 0$.

Observe first that, by the assumptions on A and B , the matrix rings $\mathbb{F}[A]$ and $\mathbb{F}[B]$ are fields. Moreover, all eigenvalue $\alpha \in \mathcal{E}_A$ and $\beta \in \mathcal{E}_B$ have \mathbb{F} -linearly independent powers up to degree $m-1$ and, respectively, $n-1$, being $\mathbb{F}(\alpha) \cong \mathbb{F}[A]$ and $\mathbb{F}(\beta) \cong \mathbb{F}[B]$, which are Galois extensions of \mathbb{F} of degree m and, respectively, n .

By a classical result on Kronecker products (see, e.g., [2, Theorem 1, p. 411] for $\mathbb{F} = \mathbb{R}$, whose generalization to finite fields is straightforward) the set of eigenvalues of M is

$$\mathcal{E}_M = \left\{ \sum_{\substack{0 \leq i < m \\ 0 \leq j < n}} z_{i,j} \alpha^i \beta^j : \alpha \in \mathcal{E}_A, \beta \in \mathcal{E}_B \right\}, \quad (18)$$

where all eigenvalues are considered as elements in some common field extension.

So, $\ker \psi = \{0\}$ if and only if each sum in (18) is nonzero. In other words, for any two $\alpha \in \mathcal{E}_A$ and $\beta \in \mathcal{E}_B$, the products $\{\alpha^i \beta^j\}_{i < m, j < n}$ are \mathbb{F} -linearly independent. By [8, Proposition 5.1 and Theorem 5.5], this condition is equivalent to

$$\mathbb{F}(\alpha) \cap \mathbb{F}(\beta) = \mathbb{F}.$$

Since the intersection of $\mathbb{F}(\alpha)$ and $\mathbb{F}(\beta)$ is the field extension of \mathbb{F} of degree $\gcd(m, n)$ (see [9, Theorem 2.6]), the proof is concluded. \square

5. The cardinality of subsets of $\mathbb{F}[A]S\mathbb{F}[B]$

In this section we will explicitly compute the cardinality of the set $\mathbb{F}[A]S\mathbb{F}[B]$ whose relevance in Cryptography is discussed in [1, 10]. Define the space of polynomials

$$\mathcal{P}^k[s] = \{p(s) \in \mathbb{F}[s] : \deg p < k\}, \quad k = 0, 1, \dots$$

being, for instance, $\mathcal{P}^0 = \{0\}$ and $\mathcal{P}^1 = \mathbb{F}$.

Note that, given a square matrix M with $d = \deg \mu_M$,

$$\mathcal{P}^0[M] \subset \mathcal{P}^1[M] \subset \dots \subset \mathcal{P}^{d-1}[M] \subset \mathcal{P}^d[M] = \mathcal{P}^k[M], \quad \forall k \geq d.$$

The main objective of this section consists in calculating the cardinality of the set

$$\mathcal{M}_{A,B;S}^{h,k} = \mathcal{P}^h[A]S\mathcal{P}^k[B] \subseteq \mathbb{F}^{m \times n}.$$

Theorem 5.1. *Let $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, and $S \in \mathbb{F}^{m \times n}$ such that $\mathcal{V}_{A,B;S} = \mathbb{F}^{m \times n}$. Then, for any $0 \leq h \leq m$ and $0 \leq k \leq n$,*

$$|\mathcal{M}_{A,B;S}^{h,k}| = \frac{(q^h - 1)(q^k - 1)}{q - 1} + 1.$$

In order to demonstrate this statement, some specific notation and one preparatory lemma are needed.

First, for every $h \leq m$, let

$$\mathbb{F}^{h;m} = \{x \in \mathbb{F}^m : x_i = 0, \forall i = h, \dots, m-1\},$$

being therefore $\mathbb{F}^h \cong \mathbb{F}^{h;m} \subseteq \mathbb{F}^m$. Define, for every $h \leq m$ and $k \leq n$, the bilinear map

$$\begin{aligned} \varphi^{h,k} : \mathbb{F}^{h;m} \times \mathbb{F}^{k;n} &\rightarrow \mathbb{F}^{m \times n} \\ (x, y) &\mapsto xy^\top \end{aligned} \quad (19)$$

and, for the sake of simplicity, denote its image by

$$\Phi^{h,k} = \varphi^{h,k}(\mathbb{F}^{h;m} \times \mathbb{F}^{k;n}). \quad (20)$$

Lemma 5.2. *Let A , B , and S as in Theorem 5.1. Then $|\mathcal{M}_{A,B;S}^{h,k}| = |\Phi^{h,k}|$.*

PROOF. Consider the map ψ defined in (17). We claim that $\psi(\Phi^{h,k}) = \mathcal{M}_{A,B;S}^{h,k}$. Actually, for every $M \in \mathcal{M}_{A,B;S}^{h,k}$, there exist $(x, y) \in \mathbb{F}^{h;m} \times \mathbb{F}^{k;n} \subseteq \mathbb{F}^m \times \mathbb{F}^n$ such that

$$M = \left(\sum_{0 \leq i < h} x_i A^i \right) S \left(\sum_{0 \leq j < k} y_j B^j \right) = \sum_{\substack{0 \leq i < m \\ 0 \leq j < n}} x_i y_j A^i S B^j = \psi(xy^\top) \in \psi(\Phi^{h,k}).$$

Therefore, $|\mathcal{M}_{A,B;S}^{h,k}| \leq |\Phi^{h,k}|$. Moreover, when $\mathcal{V}_{A,B;S} = \mathbb{F}^{m \times n}$, ψ is injective and so $\Phi^{h,k} \leftrightarrow \mathcal{M}_{A,B;S}^{h,k}$. \square

Observe that this lemma shows that the cardinality of $\mathcal{M}_{A,B;S}^{h,k}$ is independent of the choice of A , B , and S when condition (3) is met.

The problem is now reduced to the computation of the cardinality of $\Phi^{h,k}$, defined in (20).

PROOF (OF THEOREM 5.1). Consider again the map $\varphi^{h,k}$, defined in (19), and observe that

$$\mathbb{F}^{h;m} \times \mathbb{F}^{k;n} = (\varphi^{h,k})^{-1}(\Phi^{h,k}) = \bigcup_{Z \in \Phi^{h,k}} (\varphi^{h,k})^{-1}(Z).$$

Consequently, since the inverse images are disjoint,

$$q^h q^k = |\mathbb{F}^{h;m} \times \mathbb{F}^{k;n}| = \left| \bigcup_{Z \in \Phi^{h,k}} (\varphi^{h,k})^{-1}(Z) \right| = \sum_{Z \in \Phi^{h,k}} |(\varphi^{h,k})^{-1}(Z)|.$$

To compute the value of the summation, we have to consider two situations.

- When $Z = 0$, $\varphi(x, y) = xy^\top = 0$ if and only if all the products of each component of x and each component of y are zero if and only if $x = 0$ and $y = 0$ (1 case), $x = 0$ and $y \neq 0$ ($q^k - 1$ cases), or $x \neq 0$ and $y = 0$ ($q^h - 1$ cases). Therefore, $|\varphi^{-1}(0)| = q^h + q^k - 1$.
- If $Z \neq 0$, observe that, by the bilinearity of $\varphi^{h,k}$, $\varphi^{h,k}(x, y) = \varphi^{h,k}(\alpha x, \alpha^{-1}y)$ for every $\alpha \in \mathbb{F} \setminus \{0\}$.

On the other hand, if $\varphi^{h,k}(x, y) = \varphi^{h,k}(\tilde{x}, \tilde{y})$ then $\tilde{x} = \alpha x$ and $\tilde{y} = \alpha^{-1}y$ for some $\alpha \neq 0$. Indeed, considering only the indexes i and j such that $x_i y_j = \tilde{x}_i \tilde{y}_j \neq 0$, we get that

$$\frac{x_i}{\tilde{x}_i} = \frac{\tilde{y}_j}{y_j}.$$

By the independency of the indices, it follows that $\alpha = \frac{x_i}{\tilde{x}_i} = \frac{\tilde{y}_j}{y_j}$ for every i, j . So, we conclude that $|(\varphi^{h,k})^{-1}(Z)| = |\mathbb{F} \setminus \{0\}| = q - 1$.

Putting all together,

$$\begin{aligned} q^h q^k &= |(\varphi^{h,k})^{-1}(0)| + \sum_{Z \in \Phi^{h,k} \setminus \{0\}} |(\varphi^{h,k})^{-1}(Z)| \\ &= q^h + q^k - 1 + \sum_{Z \in \Phi^{h,k} \setminus \{0\}} (q - 1) = q^h + q^k - 1 + (|\Phi^{h,k}| - 1)(q - 1), \end{aligned}$$

whence

$$|\Phi^{h,k}| = \frac{q^h q^k - q^h - q^k + 1}{q - 1} + 1 = \frac{(q^h - 1)(q^k - 1)}{q - 1} + 1.$$

Finally, the claim follows by Lemma 5.2. \square

References

- [1] M.-C. Chang, On a matrix product question in cryptography, Linear Algebra Appl. 439 (7) (2013) 1742–1748. [doi:10.1016/j.laa.2013.05.013](https://doi.org/10.1016/j.laa.2013.05.013).
- [2] P. Lancaster, M. Tismenetsky, The theory of matrices : with applications, Computer science and applied mathematics, Academic Press, Orlando, 1985.
- [3] G. Micheli, Cryptanalysis of a non-commutative key exchange protocol (2013). [arXiv:1306.5326](https://arxiv.org/abs/1306.5326).

- [4] M. L. J. Hautus, Controllability and observability conditions of linear autonomous systems, *Nederl. Akad. Wetensch. Proc. Ser. A* 72 (1969) 443–448.
- [5] D. Shemesh, Common eigenvectors of two matrices, *Linear Algebra Appl.* 62 (1984) 11–18. [doi:10.1016/0024-3795\(84\)90085-5](https://doi.org/10.1016/0024-3795(84)90085-5).
- [6] T. J. Laffey, Simultaneous reduction of sets of matrices under similarity, *Linear Algebra Appl.* 84 (0) (1986) 123–138. [doi:10.1016/0024-3795\(86\)90311-3](https://doi.org/10.1016/0024-3795(86)90311-3).
- [7] H. Aslaksen, A. B. Sletsjøe, Generators of matrix algebras in dimension 2 and 3, *Linear Algebra Appl.* 430 (1) (2009) 1–6. [doi:10.1016/j.laa.2006.05.022](https://doi.org/10.1016/j.laa.2006.05.022).
- [8] P. M. Cohn, *Algebra*, Volume 3, 2nd Edition, John Wiley & Sons, Chichester, 1991.
- [9] R. Lidl, H. Niederreiter, *Finite Fields*, 2nd Edition, Vol. 20 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, 1997.
- [10] G. Maze, C. Monico, J. Rosenthal, Public key cryptography based on semi-group actions 1 (4) (2007) 489–507. [doi:10.3934/amc.2007.1.489](https://doi.org/10.3934/amc.2007.1.489).